THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Biomarkers in search of Precision Medicine in IBD

Citation for published version:
BOYAPATI, RAYKIRAN, Kalla, R, Satsangi, J & Ho, G-T 2016, 'Biomarkers in search of Precision Medicine in IBD', *The American Journal of Gastroenterology*, vol. 111, no. 12. https://doi.org/10.1038/ajg.2016.441

Digital Object Identifier (DOI):
10.1038/ajg.2016.441

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
The American Journal of Gastroenterology

Publisher Rights Statement:
Author's peer-reviewed manuscript as accepted for publication'

General rights
Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 07. May. 2023

THE AMERICAN JOURNAL OF Gastroenterology

Biomarkers in search of Precision Medicine in IBD

Journal:	*American Journal of Gastroenterology*
Manuscript ID	AJG-16-0758
Manuscript Type:	Clinical Review
Keywords:	Inflammatory bowel disease, Biomarkers, Crohns disease, Colitis, genetics, drug therapy, pharmacology, precision medicine
Manuscript Section:	Inflammatory bowel disease, Outcomes

SCHOLARONE™ Manuscripts

For Peer Review

Biomarkers in search of Precision Medicine in IBD

Ray K Boyapati[1,2], Rahul Kalla[2], Jack Satsangi[2], Gwo-tzer Ho[1,2]

Corresponding author: Dr Gwo-Tzer Ho

[1]MRC Centre for Inflammation Research

Queens Medical Research Institute

University of Edinburgh

47 Little France Crescent

Edinburgh, EH16 4TJ, UK

Email: gho@ed.ac.uk

Tel No: 0131-242 6653

[2]Gastrointestinal Unit

Institute of Genetics and Molecular Medicine

Western General Hospital

University of Edinburgh

Edinburgh, EH4 2XU, UK

For Peer Review

Abstract

The completion of the human genome project in 2003 represented a major scientific landmark, ushering in a new era with hopes and expectations of fresh insights into disease mechanisms and treatments. In IBD, many important discoveries soon followed, notably the identification of more than 200 genetic susceptibility loci and characterisation of the gut microbiome (1). As 'big data', driven by advances in technology, becomes increasingly available and affordable, individuals with IBD and clinicians alike yearn for tangible outcomes from the promise of 'precision medicine' – precise diagnosis, monitoring and treatment. Here we provide a commentary on the prospects and challenges of precision medicine and biomarkers in IBD. We focus on the three key areas where precision IBD will have the most impact: (1) disease susceptibility, activity and behaviour; (2) prediction of drug response and adverse effects; (3) identification of subphenotypic mechanisms to facilitate drug discovery and selection of new treatments in IBD.

For Peer Review

Precision medicine in IBD

Precision medicine is a major priority in health care, now recognised by all major stakeholders including governments, the pharmaceutical industry, clinicians and patients. In January 2015, United States President Barack Obama announced the Precision Medicine Initiative® (PMI): a concerted effort by multiple government agencies and backed by $215 million in federal funds to help facilitate a greater understanding of individual disease variability and its clinical translation (2). A major component is the PMI Cohort Program, an ambitious plan to build a national research cohort of more than one million participants in a participant-centred, data-driven framework with integrated multi-omic profiling. The PMI working group's report (3) identified key scientific opportunities including a number relating to the importance of biomarkers (**Box 1**). Similarly, the Chinese government has plans to invest 20 billion yuan (around US $3 billion) to support precision medicine research by 2030. Here in the United Kingdom (UK), the 100,000 genomes project was launched in 2012 with the goal of large-scale integration of genetic information and health records from the National Health Service (http://www.genomicsengland.co.uk). In the same year, the National Phenome Centre was launched in the UK, offering broad access to exploratory and targeted high-throughput metabolic phenotyping and computational biology facilities. These massive undertakings are game changers in the field of biomarker discovery and validation.

In IBD, successful international partnerships in genetics and microbiome research already provide grounds for realistic optimism. A major concern remains the wide-ranging nature of the stochastic elements of IBD, which represent formidable hurdles with respect to study design and measurable outcomes. A recent study published in *Cell* (4) provides a concrete

conceptual framework on how unbiased, data-driven development of personalized medicine approaches may be applicable to IBD. In this proof-of-concept study, post-prandial glucose responses to 46,898 meals were measured in 800 patients using continuous blood glucose monitoring. Combining patient-entered data with clinical and microbiota profiles, a computational algorithm was developed that accurately predicted personalized glycaemic response in a separate cohort. Notably, this algorithm was used in a follow-up dietary intervention study which yielded significantly lower post-prandial glucose levels (4). In our field, current creative research approaches (discussed in detail later) are now beginning to integrate across molecular datasets (e.g. genetic + microbiome), override traditional boundaries of disease classifications (UC vs. CD), target previously underexplored biological systems (e.g. virome and endogenous DAMPs) and most notably, increasingly rely on patient input using new technological applications to characterise the 'exposome' in IBD (5)(6)(7). Hence, a new theme of recombinant innovation is emerging with synergy arising from novel ideas within established and fresh datasets.

IBD biomarkers in clinical practice - the story so far

Numerous potential genetic, blood-based, faecal, microbial and immunological biomarkers have been proposed in IBD (8) (**Table 1**). This has recently been extensively reviewed (9)(10)(11)(12). However, apart from a few notable exceptions, biomarkers have not yet found widespread clinical application in IBD practice for a variety reasons (**Box 2**). We highlight three examples of 'biomarkers' that have roles in clinical practice, namely faecal calprotectin (FC), anti-TNF antibodies and thiopurine methyltransferase (TPMT) activity measurements. FC is a screening tool for gut inflammation (13) and to measure disease activity in IBD (14). More recently, the potential to use FC in innovative ways has been explored including as a predictive tool (e.g. to identify disease recurrence in post-operative CD (15)) and as a secondary end-point in IBD clinical trials. Detection of anti-TNF antibodies allows for expedient switching to an alternative drug (16) and avoids conventional dose escalation which is often futile, expensive and potentially hazardous (17). A recent randomized controlled trial in the setting of secondary loss of response to infliximab compared conventional dose intensification with an algorithm based approach based on serum infliximab levels and antibodies (18). Here, management dictated by drug levels and antibodies was found to be cost effective with no reduction in clinical efficacy. TPMT measurement can screen for those who are likely to experience life threatening leukopenia from thiopurines (19) and those who would benefit from a reduced initial dose. Although of some clinical benefit, these currently available examples provide some perspective to the

For Peer Review

enthusiasm towards Precision Medicine, highlighting the wide gulf between what clinicians currently have at their disposal and the ambitious aspirations for the near future.

IBD big data: Sink or swim?

The critical question is: are we on the cusp of a therapeutic revolution underpinned by the inexorable wave of big data, or will we end up drowning in a sea of potential biomarkers that we cannot translate into clinical practice? A number of critical enablers allow for optimism (**Figure 1**). First, government and industry interest and investment will continue to improve the development of large-scale prospective cohorts. Ambitious biobank projects such as the recently launched UK-wide IBD BioResource (www.ibdbioresource.nihr.ac.uk) and the aforementioned PMI Cohort Program will help overcome this first obstacle for novel IBD biomarker discovery and validation. Second, we expect advances in high throughput technology to allow for quicker, cheaper and more efficient testing of large biobanks. Costs for DNA sequencing have shrunk by more than 10 million fold since 1998 – something almost inconceivable at the time. It is similarly difficult to comprehend what might be possible in the future. In concert, advances in computational power will continue to facilitate production and analysis of massive amounts of accrued IBD –omics data. Third, we foresee the influence of the exposome being clarified from an increasing emphasis on patient inputted data. Technological advances such as personal mobile devices for real-time monitoring and electronic health record integration will provide a platform for prospective and progressive data collection. In IBD, there is increasing use of real-time feedback of clinical information, environmental factors and disease activity (e.g. home FC kits) back to clinicians and researchers (20).

These advancements require even closer levels of cooperation and collaboration between researchers. Furthermore, there is a greater need for more creative analytical approaches that will involve a model of continuous learning and analysis. Centres of excellence dedicated to Precision Medicine and Big Data analyses are now currently being set up globally. Some health systems may be better equipped to provide the long term biological data and clinical follow-up that captures the disease population in its widest sense (e.g. arguably, the UK National Health System).

Notwithstanding all these major interventions, there is a need for a dose of realism. In cancer research, where investment has been far greater, there has been a decrease in the number of FDA approved protein biomarkers over the last decade (21). In IBD, it is notable that biomarkers in existing use such as faecal calprotectin were found through hypothesis based investigation (22) rather than high throughput methods or *in silico* database analysis.

For Peer Review

Presently, discovery-based approaches are burdened by a number of significant challenges (summarized in **Box 3**). In particular, selection bias from convenience sampling and data overfitting can result in over-interpretation of 'significant' *p*-values, potentially wasting valuable resources on random noise. For example, a host of studies have identified genetic polymorphisms as predictors of therapeutic response in IBD (9) but these have not been consistently replicated. There remains room for advancements based on discoveries in related inflammatory conditions, serendipity and organic scientific thinking, although big data now forms the ground for the generation of new hypotheses.

Precision Medicine in IBD: Recent progress

Multiple lines of evidence show progress towards Precision Medicine and we provide a brief overview of several promising studies in three key areas in IBD (disease susceptibility, activity and behaviour; prediction of drug response and adverse effects; and identification of subphenotypic mechanisms for drug selection and discovery). Collectively, they highlight a few recurrent themes. First, the enormous potential of how big data can be even more powerful when different data platforms are integrated. Second, the process involved in validating these increasingly complex findings remains difficult, as does the path for biomarkers to reach clinical application. In this context, it seems likely that the future maybe based on computational predictive modelling, incorporating many biomarkers (as shown by Zeevi et al), and evident in artificial intelligence systems that permeate current daily living (e.g. weather forecasting and voice recognition search systems). This challenges the traditional criteria for a good biomarker being simple, accurate, easy to perform, minimally invasive, cheap, rapid and reproducible. Hence, some re-orientation of IBD patients' and clinicians' perspectives may be needed.

Disease susceptibility, activity and behaviour

Here genetic studies lead the way with the latest meta-analysis now involving 50 000 IBD individuals implicating more than 200 susceptibility loci (23)(24). This information provided hitherto unknown insights into disease mechanisms and biological pathways such as autophagy and IL23/Th17. The role for genetic data in predicting susceptibility, activity and disease behaviour is however less strong. For example, the strongest genetic signal, NOD2 status has been associated with ileal and fibrostenosing disease but carriage of the NOD2 mutant allele is uncommon. Genetic information allied with other biological data (e.g. pheno-genomic status), maybe more informative. A combination of clinical, serological and genotypic data has been used to help predict the risks of surgery in CD (25). Recently, a study of 29,838 IBD patients which tested for genetic-phenotype associations found that

For Peer Review

predictive models based on generated genetic risk scores strongly distinguished colonic from ileal CD (26).

Beyond genetics, epigenetics is emerging as a further tier of information that could complement genome wide association studies (GWAS) (27). Several epigenome wide studies have been published in IBD and other diseases (28–32). These studies identify epigenetic mechanisms as a potential interface between genetics and disease. Highly significant enrichment of methylation changes have been shown to occur around GWAS single nucleotide polymorphisms, in particular the HLA region and MIR21 (28). The Roadmap Epigenomics Project recently published the most comprehensive epigenomic reference providing information on 111 new reference epigenomes and how these control gene expression in humans (33). Equally, other epigenetic mechanisms such as miRNAs are being explored for their roles in IBD pathogenesis (34). These short strands of non-coding RNA (~22nt long) have been shown to regulate key GWAS pathways in IBD including autophagy and epithelial barrier integrity (34).

The microbiome or 'other genome' – the collective genome of the gut microbiota – represents a further giant dimension in big data in IBD and other complex multifactorial conditions such as diabetes and obesity (4,35–37). The NIH-funded Human Microbiome Project Consortium has generated in-depth data on the genomic information and microbial community structure and function in humans. Some early studies show potential for clinical translation towards Precision Medicine. In the largest study of treatment naïve CD patients to date, Gevers and colleagues found ileal microbiome signatures were predictive of CD and observed this even in the absence of overt inflammation (38). In one recent prospective study of paediatric IBD receiving anti-TNF therapy, faecal microbial diversity resembled controls in patients who responded to anti-TNF therapy versus non-responders (39). This shift in microbial diversity in responders was also seen in a paediatric UC study of steroid response (40). Furthermore, microbial populations may help predict disease course as illustrated by a study of post-operative recurrence in CD, where a decreased population of *Faecalibacterium prausnitzii* in the resected ileum correlated with increased risk of recurrence (41). We anticipate rapid progress in this area with many opportunities for the public and patients to contribute their faecal samples for analysis (e.g. the American Gut Project).

Away from predicting susceptibility and behaviour in IBD, there is an unmet need for sensitive biomarkers to measure gut mucosal inflammation to provide objective data on disease activity and guide response to treatment. Presently, C-reactive protein (CRP) and faecal calprotectin have better negative predictive values and are thus more useful in

For Peer Review

excluding significant inflammatory signals. Some progress can be expected in modalities to image inflammation (e.g. MRI) with or without the use of specific in-vivo labelling of inflammatory cells or targets. A recent study used confocal laser endomicroscopy to detect fluorescent antibody labelled membrane-bound TNF (mTNF) in intestinal immune cells of 25 CD patients (42). Patients with high numbers of mTNF+ cells were more likely to respond to adalimumab at week 12 than those with low amounts of mTNF+ cells (92% versus 15%). Furthermore, a re-thinking of ways to measure established biomarkers such as measuring serum (rather than faecal) calprotectin (43) may improve the performance and applicability of these tests. In search of better disease activity prediction, current approaches as exemplified by the EMBARK study, assess a panel of biomarkers by their correlation with endoscopy and radiological findings as the best reference measure (44). However, although such reference measures capture disease extent; location and burden well, they are not specific enough to evaluate disease behaviour, complications and progression (44,45). As will be discussed later, better biomarkers of activity may come from a refined approach measuring specific downstream effects of the biological pathway targeted (e.g. Th17-family of cytokines in IL23/IL17 inhibition).

Prediction of drug response and adverse effects

Exploiting the wealth of genetic data, the combination of phenotypic information with multiple susceptibility loci has been shown to be predictive of primary non-response in anti-TNF therapy in paediatric IBD (46). There have been some notable successes in transcriptomics (gene expression) in IBD. Lee et al showed that CD8+ T-cell immune signatures are better at predicting disease course than traditional clinical or serological markers (47). Hence, this approach is useful to select individuals that might benefit from more aggressive medical treatment. At the mucosal level, one study of infliximab in UC used gene signatures to separate responders from non-responders with 95% sensitivity and 85% specificity (48).

In terms of predicting adverse effects, a recent study found that a nonsynonymous single-nucleotide polymorphism in NUDT15 has a greater effect than TPMT variants in patients of Korean ancestry in predicting thiopurine induced leukopenia (49). GWAS of azathioprine induced pancreatitis found increased susceptibility for HLA-DQA1-HLA-DRB1 variants with a 2.5 fold increased risk in heterozygotes and a 5 fold increased risk in homozygotes at rs2647087 (50). Although an important finding, this potential biomarker highlights some of the difficulties encountered in translation to the clinic. For example, the low pre-test probability of pancreatitis means that even in the highest risk homozygotes, there is an 83% chance of taking thiopurines without developing pancreatitis. Would this justify the exclusion

For Peer Review

of thiopurines in these patients? Perhaps not, but in a future of greater therapeutic options, this may be more feasible and in the meantime, will improve risk counselling and awareness. In addition, the number needed to test in this study was 76, making it an expensive option for screening. However, cheaper point of care testing in the future, and the possible combination with other biomarkers could change the economics of such a test.

Identification of subphenotypic mechanisms to facilitate drug discovery and selection of new treatments

IBD clinicians will have increasing number of drugs available with over 20 currently in the developmental pipeline (51)(52). Rather than a sequential approach of trying one drug after another, one of the goals for Precision Medicine is to identify individuals or disease phenotypes that are better suited for a particular drug from the outset (e.g. anti-leukocyte migration versus anti-TNF). This direction is appealing and cogent where recent advances in oncology and virology have shown the way. In breast cancer, human epithelial growth factor (HER2) positivity provides prognostic information (more aggressive phenotype with higher recurrence rates) as well as therapeutic choice (response to monoclonal antibodies targeting HER2 such as trastuzumab) (53). In metastatic colorectal cancer, KRAS gene mutations predict response to anti-epidermal growth factor receptor (anti-EGFR) monoclonal antibody therapy (54). Notably, the early studies were performed based on a hypothesis developed from an understanding of EGFR biology (55); subsequent retrospective subset analysis of randomized controlled trials provided strong evidence for clinical use. In non-small cell lung cancer, mutations in multiple oncogenes including ALK and EGFR can help direct tyrosine kinase therapy (56). In hepatitis C, prior to the direct acting antiviral revolution, IL28B genotype helped predict the likelihood of sustained viral response to interferon and ribavirin therapy.

The current case of anti-TNF is instructive and highlights the difficulties ahead in developing a biomarker based decision-making model. Although therapeutic drug monitoring and antibody testing are important to inform dosing and the decision to switch biologic, there is currently no clinically useful predictor of anti-TNF response prior to initiation. Retrospective studies and *post-hoc* analysis of large clinical trials found a number of phenotypic, demographic and biochemical markers that help predict response but the effect size and evidence is not convincing enough to be clinically useful (57). At a molecular level, attempts were made to investigate anti-TNF response based on TNF-α gene polymorphisms, but these produced inconsistent results (58). Baseline TNF levels were found to be higher in those who did not respond to infliximab at 10 weeks (59), but a subsequent larger study did

For Peer Review

not find any such relationship (60). Prediction of anti-TNF response has also been attempted with *in vivo* imaging (61) and mucosal gene expression signatures (62)(48). Furthermore, there are ongoing research efforts to elucidate the factors that determine response to anti-TNF drugs, such as the multi-centre UK PANTS study in active luminal Crohn's disease. A more incisive and economical investigative model might be to start from a fresh platform, by building the study of biomarkers into clinical trials in IBD.

Incorporating biomarkers in clinical trials

The unexpected failures of secukinumab (63) and tofacitinib (64) in CD, anrukinzumab in UC (65) and abatacept in CD and UC (66) have raised the question of whether molecular based stratification with biomarkers based on underlying disease mechanisms could reveal subgroups within the traditional CD and UC groups, who would selectively benefit from certain therapies (**Figure 2**). A number of large clinical trials are seeking to incorporate biomarkers into trial design (67), with the notion that biomarker development and clinical trial design should operate hand-in-hand. Multiplex IBD biomarker panels are created at diagnosis to obtain a snapshot of a patient's molecular disease profile. These biomarker profiles can be used in *post hoc* analyses to investigate their relationship to drug response (**Figure 2**).

Incorporating new biomarkers in clinical trials however, is challenging and expensive. The cost of drug development is inversely related to the size of the target population, which is necessarily smaller in personalized therapy. Furthermore, stratification using all possible biomarker combinations is impractical given resource and trial participant limitations. At some point, decisions need to be made regarding which biomarkers are incorporated into trials. Another challenge is the standardization of classification/stratification among studies. Although more information, including molecular profiling, is generally better than less information, distinct and stable classification also has its advantages including comparability between trials, and improved communication between clinicians and researchers. Increasingly targeted therapies will require more defined biomarkers to measure their effects on the respective biological pathways. This sets the scene for stratified clinical trials as seen in oncology for example.

Conclusion: Meta-ideas in IBD

Clinicians long for a future where a newly diagnosed IBD patient can have his/her genetic, microbiome and immune profile measured at the outset; then matched to the most appropriate biologic or immunosuppressive treatment based on likelihood of response/adverse effects. These IBD individuals will be informed of what 'exposome' to

For Peer Review

modify and report on their disease activity using the set of optimal biomarkers. At all levels, one can expect a continuous feedback of new data from respective patients, which will further improve the dataset for biomarker discovery.

For some time, the barrier for real progress has been affordable access to big data. This is now within reach. The digital age now involves patients as active contributors to research data and creates an expansive network of researchers spanning across the traditional disease, biological pathway and systems' boundaries. The digitalisation of clinical data and penetration of artificial intelligence into science are entirely new dimensions in play. Notwithstanding all these factors, we believe that the force that will shape Precision Medicine in IBD will be the focus on meta-ideas - ideas that will further accelerate the production or transmission of new ideas. Established organisations such NASA increasingly rely on 'open innovation' or crowdsourcing to find new solutions to their most difficult and intractable problems. The availability of 'big data' is no longer the rate-limiting step to progress; instead, it is the clinician or researcher's ingenuity in leveraging these assets into new knowledge.

For Peer Review

FIG 1: CRITICAL ENABLERS IN THE FLOW OF PRECISION MEDICINE IN IBD

FIG 2: INCORPORATING BIOMARKERS INTO THE DESIGN OF CLINICAL TRIALS IN IBD

TABLE 1: Summary of potential and current applications of biomarkers relevant to the 3 key areas of Precision Medicine in IBD.

Disease susceptibility, activity and behaviour	Prediction of adverse effect and therapeutic response	Identification of subphenotypic mechanism
Susceptibility/behaviour Genetics and serological markers *NOD2, ATG16L1, IL23R* ASCA, ANCA Microbial Compositional microbiome shifts (Enterobacteriaceae, Enterococcaceae, Lachnospiraceae (Blautia, Dorea), Prevotella) Decreased *Faecalibacterium prausnitzii* Increased adherent invasive *E. coli* *Activity* CRP, ESR, serum albumin, stool calprotectin, lactoferrin, S100A12, Magnetic resonance imaging Composite stool calprotectin and blood MMP9 and IL-22	*Adverse effect* *TPMT* *NUDT15* *HLA-DQA1-HLA-DRB1* *Therapeutic response* Antibodies towards TNF Anti-TNF and thiopurine metabolites mTNF+ in intestinal immune cells by confocal laser endomicroscopy	Cytokines and transcriptomics IL-13, IL-23 Th17 cytokines (IL-17A, IL-17F, IL-22) IL-6, IL-1, IL-8. IL-10, TNFα CD8-T cell signatures Soluble ST2 and IL-33 Mucosal indoleamine 2,3 dioxygenase-1

For Peer Review

BOX 1

BIOMARKER FOCUSSED SCIENTIFIC OPPORTUNITIES IDENTIFIED BY THE PMI

- Biomarker discovery for identification of individuals with a higher risk of developing disease to help more rational prevention efforts
- Interrogation of the wide variation in therapeutic response and adverse reactions
- Novel classification systems which transcend the existing grouping based on symptoms, signs and laboratory results by using molecular characterisation
- Using biomarkers to assign patients into a variety of clinical trials targeting subsets based on these biomarkers to help with development of novel therapies
- Translating pre-existing environmental and genetic risk factors into conclusions on disease causes and population impact with population based cohort studies as well as identifying new associations.

BOX 2

WHY HAVE MOST POTENTIAL IBD BIOMARKERS NOT FOUND THEMSELVES IN WIDESPREAD CLINICAL USE?

- Failures on the classic qualities of an ideal biomarker (68) (simple, accurate, easy to perform, minimally invasive, cheap, rapid, reproducible)
- Unclear or uncertain clinical utility: i.e. does not provide clinically useful information upon which to make decisions
    a. low sensitivity/specificity
    b. low prognostic/predictive values
- Lack of validation in independent cohorts or have had inconsistent results when validation has been attempted.
- Some areas (such as microbiome based biomarkers) are in their infancy

BOX 3

CHALLENGES WITH –OMICS RESEARCH

- Large costs associated with biomarker validation for those biomarkers proposed by unbiased -omics testing
- Potential confounders including interaction between the different ‘omes’ (e.g. microbiome studies with effect of host genome), disease activity, duration, location and effects of drug treatment, study design, heterogeneous cohorts
- Selection bias using convenience sampling

For Peer Review

- Increased flexibility and non-linearity in algorithms leading to overfitting
- Lack of support from pharmaceutical companies not wanting to fragment markets
- Reluctant adoption and acceptance by physicians and patients
- So far, mainly cross-sectional studies i.e. one point in time rather than continuous prospective studies
- Electronic medical record integration – difficulties in standardization, poor quality and granularity of inputted data. National approach easier in some countries (e.g. UK with NHS) than others e.g. USA
- Privacy and data security
- Standardization of all steps in the process of biomarker discovery is optimal but in practice, difficult to achieve.
- Teamwork and collaboration across institutions (particularly critical in relatively low incidence diseases such as IBD) with use of standard protocols and large cohorts.

POTENTIAL SOLUTIONS

- Data-driven approaches such as network interference
- Prospective studies with multiple time points
- Standardized method of sample acquisition
- Use of homogenous patient subsets and studies of subjects with no prior medical therapy

For Peer Review

CURRENT KNOWLEDGE AND WHAT IS NEW HERE

- Despite recent progress in understanding IBD pathogenesis, there are only a few biomarkers (faecal calprotectin and C-reactive protein) that are widely used in clinical practice.
- Successful biomarker discovery is hampered by many factors particularly, the wide-ranging nature of the stochastic elements of IBD, which represent formidable barriers to study design and measurable outcomes.
- 'Big data' in IBD (encompassing including host multi-omic profile, microbiome and exposome) and increasingly powerful computational approaches are now within reach.
- Precision Medicine (PM) Initiative is a major program focused towards achieving a greater understanding of individual disease variability and its clinical translation. Together with many international stakeholders involved in PM, this impetus represents a game-changer directly relevant to IBD.
- We focus on the three key areas where precision IBD will have the most impact: (1) disease susceptibility, activity and behaviour; (2) prediction of drug response and adverse effects; (3) identification of subphenotypic mechanisms to facilitate drug discovery and selection of new treatments in IBD.
- 'Meta-ideas', the pipeline to generate of new and innovative ideas to leverage 'Big data' into tangible clinical translation now represents the most exciting challenge in biomarker discovery in IBD.

For Peer Review

References

1. Structure, function and diversity of the healthy human microbiome. Nature 2012;486:207–214.
2. Collins FS, Varmus H. A New Initiative on Precision Medicine. N. Engl. J. Med. 2015;372:793–5.
3. Hudson K, Lifton R, Patrick-Lake B. The Precision Medicine Initiative Cohort Program – Building a Research Foundation for 21st Century Medicine. Precis. Med. Initiat. Work. Gr. Rep. to Advis. Comm. to Dir. NIH 2015;103.
4. Zeevi D, Korem T, Zmora N, *et al.* Personalized Nutrition by Prediction of Glycemic Responses. Cell 2015;163:1079–1094.
5. Boyapati RK, Rossi AG, Satsangi J, *et al.* Gut mucosal DAMPs in IBD: from mechanisms to therapeutic implications. Mucosal Immunol 2016;
6. Rogler G, Vavricka S. Exposome in IBD: recent insights in environmental factors that influence the onset and course of IBD. Inflamm. Bowel Dis. 2015;21:400–408.
7. Virgin HW. The virome in mammalian physiology and disease. Cell 2014;157:142–150.
8. Louis E, Mary J, Vernier–Massouille G, *et al.* Maintenance of Remission Among Patients With Crohn's Disease on Antimetabolite Therapy After Infliximab Therapy Is Stopped. Gastroenterology 2012;142:63–70.e5.
9. Mosli MH, Sandborn WJ, Kim RB, *et al.* Toward a personalized medicine approach to the management of inflammatory bowel disease. Am. J. Gastroenterol. 2014;109:994–1004.
10. Sands BE. Biomarkers of Inflammation in Inflammatory Bowel Disease. Gastroenterology 2015;1275–1285.

For Peer Review

11. Dubinsky M, Braun J. Diagnostic and Prognostic Microbial Biomarkers in Inflammatory Bowel Diseases. Gastroenterology 2015;149:1265–1274.e3.

12. McGovern DPB, Kugathasan S, Cho JH. Genetics of Inflammatory Bowel Diseases. Gastroenterology 2015;149:1163–1176.e2.

13. Rheenen PF van, Vijver E Van de, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. BMJ 2010;341:c3369.

14. Lin J-F, Chen J-M, Zuo J-H, *et al.* Meta-analysis: fecal calprotectin for assessment of inflammatory bowel disease activity. Inflamm. Bowel Dis. 2014;20:1407–15.

15. Wright EK, Cruz P De, Gearry R, *et al.* Fecal Biomarkers in the Diagnosis and Monitoring of Crohn's Disease. Inflamm. Bowel Dis. 2014;0:1–10.

16. Nanda KS, Cheifetz AS, Moss AC. Impact of antibodies to infliximab on clinical outcomes and serum infliximab levels in patients with inflammatory bowel disease (IBD): a meta-analysis. Am. J. Gastroenterol. 2013;108:40–7; quiz 48.

17. Steenholdt C, Brynskov J, Thomsen OØ, *et al.* Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. Gut 2014;63:919–27.

18. Steenholdt C, Brynskov J, Thomsen OO, *et al.* Individualised therapy is more cost-effective than dose intensification in patients with Crohn's disease who lose response to anti-TNF treatment: a randomised, controlled trial. Gut 2014;63:919–927.

19. Lennard L. TPMT in the treatment of Crohn's disease with azathioprine. Gut 2002;51:143–146.

20. Deen WK Van, Meulen-de Jong AE van der, Parekh NK, *et al.* Development and Validation of an Inflammatory Bowel Diseases Monitoring Index for Use With Mobile Health Technologies. Clin. Gastroenterol. Hepatol. 2016;

21. Pavlou MP, Diamandis EP, Blasutig IM. The long journey of cancer biomarkers from

For Peer Review

the bench to the clinic. Clin. Chem. 2013;59:147–157.

22. Roseth AG, Fagerhol MK, Aadland E, *et al.* Assessment of the neutrophil dominating protein calprotectin in feces. A methodologic study. Scand. J. Gastroenterol. 1992;27:793–798.

23. Liu JZ, Sommeren S van, Huang H, *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat. Genet. 2015;47:979–989.

24. Jostins L, Ripke S, Weersma RK, *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.

25. Dubinsky MC, Kugathasan S, Kwon S, *et al.* Multidimensional prognostic risk assessment identifies association between IL12B variation and surgery in Crohn's disease. Inflamm. Bowel Dis. 2013;19:1662–70.

26. Cleynen I, Boucher G, Jostins L, *et al.* Inherited determinants of Crohn ' s disease and ulcerative colitis phenotypes : a genetic association study. Lancet 2015;6736:1–12.

27. Ventham NT, Kennedy N a, Nimmo ER, *et al.* Beyond gene discovery in inflammatory bowel disease: the emerging role of epigenetics. Gastroenterology 2013;145:293–308.

28. Adams AT, Kennedy NA, Hansen R, *et al.* Two-stage Genome-wide Methylation Profiling in Childhood-onset Crohn's Disease Implicates Epigenetic Alterations at the VMP1/MIR21 and HLA Loci. Inflamm. Bowel Dis. 2014;

29. Nimmo ER, Prendergast JG, Aldhous MC, *et al.* Genome-wide methylation profiling in Crohn's disease identifies altered epigenetic regulation of key host defense mechanisms including the Th17 pathway. Inflamm. Bowel Dis. 2012;18:889–899.

30. McDermott E, Ryan EJ, Tosetto M, *et al.* DNA Methylation Profiling in Inflammatory Bowel Disease Provides New Insights into Disease Pathogenesis. J. Crohns. Colitis 2016;10:77–86.

For Peer Review

31. Yuan W, Xia Y, Bell CG, *et al.* An integrated epigenomic analysis for type 2 diabetes susceptibility loci in monozygotic twins. Nat. Commun. 2014;5:5719.

32. Dick KJ, Nelson CP, Tsaprouni L, *et al.* DNA methylation and body-mass index: a genome-wide analysis. Lancet (London, England) 2014;383:1990–8.

33. Romanoski CE, Glass CK, Stunnenberg HG, *et al.* Epigenomics: Roadmap for regulation. Nature 2015;518:314–316.

34. Kalla R, Ventham NT, Kennedy NA, *et al.* MicroRNAs: new players in IBD. Gut 2015;64:504–17.

35. Jostins L, Ripke S, Weersma RK, *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 2012;491:119–24.

36. Chatelier E Le, Nielsen T, Qin J, *et al.* Richness of human gut microbiome correlates with metabolic markers. Nature 2013;500:541–6.

37. Qin J, Li Y, Cai Z, *et al.* A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature 2012;490:55–60.

38. Gevers D, Kugathasan S, Denson L a, *et al.* The treatment-naive microbiome in new-onset Crohn's disease. Cell Host Microbe 2014;15:382–92.

39. Kolho K-L, Korpela K, Jaakkola T, *et al.* Fecal Microbiota in Pediatric Inflammatory Bowel Disease and Its Relation to Inflammation. Am. J. Gastroenterol. 2015;921–930.

40. Michail S, Durbin M, Turner D, *et al.* Alterations in the gut microbiome of children with severe ulcerative colitis. Inflamm. Bowel Dis. 2012;18:1799–1808.

41. Sokol H, Pigneur B, Watterlot L, *et al.* Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. U. S. A. 2008;105:16731–6.

42. Atreya R, Neumann H, Neufert C, *et al.* In vivo imaging using fluorescent antibodies to tumor necrosis factor predicts therapeutic response in Crohn's disease. Nat. Med.

For Peer Review

2014;20:313–8.

43. Meuwis M-A, Vernier-Massouille G, Grimaud JC, *et al.* Serum calprotectin as a biomarker for Crohn's disease. J. Crohn's Colitis 2013;7:e678–e683.

44. Faubion WAJ, Fletcher JG, O'Byrne S, *et al.* EMerging BiomARKers in Inflammatory Bowel Disease (EMBARK) study identifies fecal calprotectin, serum MMP9, and serum IL-22 as a novel combination of biomarkers for Crohn's disease activity: role of cross-sectional imaging. Am. J. Gastroenterol. 2013;108:1891–1900.

45. Faubion WA, Fletcher JG, Villiers WJ de, *et al.* Mo1680 Assessing Crohn's Disease Inflammatory Biomarker Diagnostic Accuracy Using Ileocolonoscopy or a Combined Ileocolonoscopy-CTE Score in the Embark Study. Gastroenterology 2012;142:S–658.

46. Dubinsky MC, Mei L, Friedman M, *et al.* Genome wide association (GWA) predictors of anti-TNFalpha therapeutic responsiveness in pediatric inflammatory bowel disease. Inflamm. Bowel Dis. 2010;16:1357–66.

47. Lee JC, Lyons PA, Mckinney EF, *et al.* Gene expression profiling of CD8 + T cells predicts prognosis in patients with Crohn disease and ulcerative colitis. 2011;121

48. Arijs I, Li K, Toedter G, *et al.* Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. Gut 2009;58:1612–1619.

49. Yang S-K, Hong M, Baek J, *et al.* A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. Nat. Genet. 2014;46:1017–1020.

50. Heap G a, Weedon MN, Bewshea CM, *et al.* HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants. Nat. Genet. 2014;46:1131–4.

51. D'Haens GR, Sartor RB, Silverberg MS, *et al.* Future directions in inflammatory bowel disease management. J. Crohns. Colitis 2014;8:726–734.

52. Danese S. New therapies for inflammatory bowel disease: from the bench to the bedside. Gut 2012;61:918–32.

For Peer Review

53. Arteaga CL, Sliwkowski MX, Osborne CK, *et al.* Treatment of HER2-positive breast cancer: current status and future perspectives. Nat. Rev. Clin. Oncol. 2011;9:16–32.

54. Allegra CJ, Jessup JM, Somerfield MR, *et al.* American Society of Clinical Oncology Provisional Clinical Opinion: Testing for KRAS Gene Mutations in Patients With Metastatic Colorectal Carcinoma to Predict Response to Anti–Epidermal Growth Factor Receptor Monoclonal Antibody Therapy. J. Clin. Oncol. 2009;27 :2091–2096.

55. Lièvre A, Bachet J-B, Corre D Le, *et al.* KRAS Mutation Status Is Predictive of Response to Cetuximab Therapy in Colorectal Cancer. Cancer Res. 2006;66 :3992–3995.

56. Lindeman NI, Cagle PT, Beasley MB, *et al.* Molecular Testing Guideline for Selection of Lung Cancer Patients for EGFR and ALK Tyrosine Kinase Inhibitors. J. Thorac. Oncol. 2013;8:823–859.

57. Siegel CA, Melmed GY. Predicting Response to Anti-TNF Agents for the Treatment of Crohn's Disease. Therap. Adv. Gastroenterol. 2009;2:245–251.

58. Chaudhary R, Ghosh S. Prediction of response to infliximab in Crohn's disease. Dig. liver Dis. 2005;37:559–563.

59. Martínez-Borra J, López-Larrea C, González S, *et al.* High serum tumor necrosis factor-alpha levels are associated with lack of response to infliximab in fistulizing Crohn's disease. Am. J. Gastroenterol. 2002;97:2350–2356.

60. Louis E, Vermeire S, Rutgeerts P, *et al.* A positive response to infliximab in Crohn disease: association with a higher systemic inflammation before treatment but not with -308 TNF gene polymorphism. Scand. J. Gastroenterol. 2002;37:818–824.

61. Brande JM Van den, Koehler TC, Zelinkova Z, *et al.* Prediction of antitumour necrosis factor clinical efficacy by real-time visualisation of apoptosis in patients with Crohn's disease. Gut 2007;56:509–517.

62. Halloran B, Chang J, Shih DQ, *et al.* Molecular patterns in human ulcerative colitis

For Peer Review

and correlation with response to infliximab. Inflamm. Bowel Dis. 2014;20:2353–63.

63. Hueber W, Sands BE, Lewitzky S, *et al.* Secukinumab, a human anti-IL-17A monoclonal antibody, for moderate to severe Crohn's disease: unexpected results of a randomised, double-blind placebo-controlled trial. Gut 2012;61:1693–700.

64. Sandborn WJ, Ghosh S, Panes J, *et al.* A Phase 2 Study of Tofacitinib, an Oral Janus Kinase Inhibitor, in Patients with Crohn's Disease. Clin. Gastroenterol. Hepatol. 2014;1485–1493.

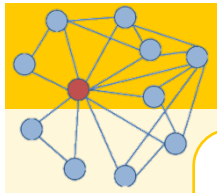
65. Reinisch W, Panés J, Khurana S, *et al.* Anrukinzumab, an anti-interleukin 13 monoclonal antibody, in active UC: efficacy and safety from a phase IIa randomised multicentre study. Gut 2015;64 :894–900.

66. Sandborn WJ, Colombel J-F, Sands BE, *et al.* Abatacept for Crohn's disease and ulcerative colitis. Gastroenterology 2012;143:62–69.e4.

67. Dancey JE, Dobbin KK, Groshen S, *et al.* Guidelines for the development and incorporation of biomarker studies in early clinical trials of novel agents. Clin. Cancer Res. 2010;16:1745–1755.

68. Vermeire S, Assche G Van, Rutgeerts P. Laboratory markers in IBD: useful, magic, or unnecessary toys? Gut 2006;55:426–31.

For Peer Review



Data Sources

Data in

Genetics, Epigenetics
Specific - ATG16L1, NOD2, IL23R
General – Genetic/epigenetic signatures

Exposome
Smoking
Diet – sugar, fat
Geography

Clinical data

Microbiome
General – microbial shifts
Specific – F Prau, AEICs

Inflammatory profile
Blood, mucosal
TNFa, IL17, IL13

Transcriptome
CD8-T cell signatures

Analysis

Process

New analytical approaches – learning algorithms

Integration of datasets

Digitalisation of clinical data

Molecular vs. clinical stratification

Precision Medicine

Outcomes

Disease susceptibility
Disease activity
Disease behaviour

Therapeutic response
Adverse effects
Prognosis

Stratification by mechanism

New Enablers

Large scale prospective cohorts

Patient inputted data

Advances in computational power

Meta-ideas and innovative access to new ideas

Incorporation into clinical trials

Fig 1

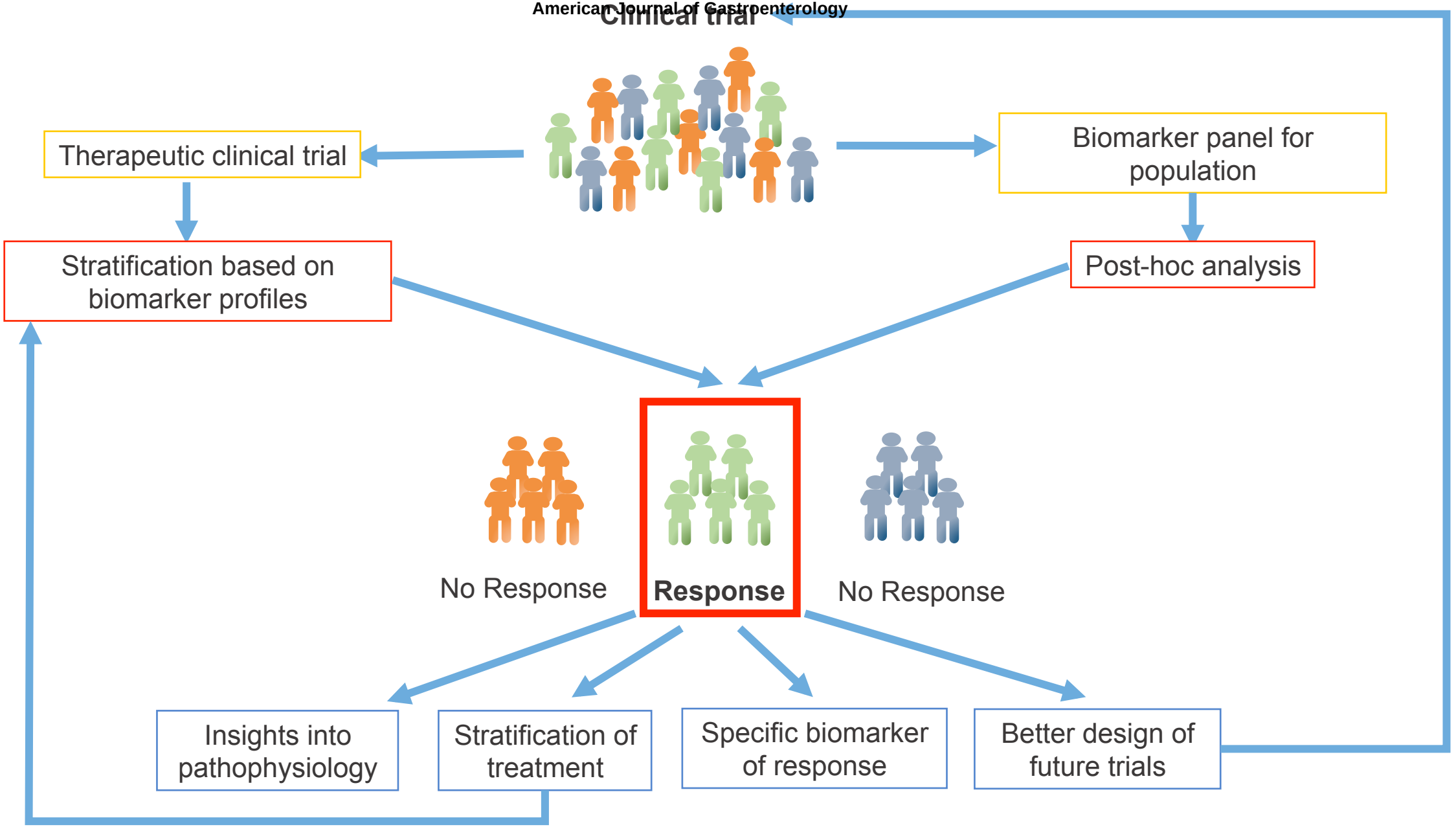

Clinical trial
Therapeutic clinical trial
Biomarker panel for population
Stratification based on biomarker profiles
Post-hoc analysis
No Response
Response
No Response
Insights into pathophysiology
Stratification of treatment
Specific biomarker of response
Better design of future trials

Fig 2